Communications on Stochastic Analysis

Volume 13 | Number 2 Article 2

6-2019

Smoothing Parameters for Recursive Kernel Density Estimators under Censoring

Yousri Slaoui

Univ. Poitiers, Lab. Math. et Appl., Futuroscope Chasseneuil, France, Yousri.Slaoui@math.univ-poitiers.fr

Recommended Citation

Slaoui, Yousri (2019) "Smoothing Parameters for Recursive Kernel Density Estimators under Censoring," *Communications on Stochastic Analysis*: Vol. 13 : No. 2 , Article 2.

DOI: 10.31390/cosa.13.2.02

Available at: https://digitalcommons.lsu.edu/cosa/vol13/iss2/2

Communications on Stochastic Analysis
Vol. 13, No. 2 (2019) Article 2 (25 pages)
DOI: 10.31390/cosa.13.2.02

SMOOTHING PARAMETERS FOR RECURSIVE KERNEL DENSITY ESTIMATORS UNDER CENSORING

YOUSRI SLAOUI*

ABSTRACT. In this paper, we are concerned with the nonparametric estimation of an unknown density under censoring. Firstly, we propose a recursive kernel density estimators under censoring, based on a stochastic approximation algorithm. Then, we showed that our recursive estimator is consistent and asymptotically normally distributed. Moreover, we describe and investigate a data-driven bandwidth selection procedure based on normal pilot bandwidth reference distributions. We showed that the proposed recursive estimators can be better than the non-recursive in terms of estimation error and much better in terms of computational costs. We corroborated these theoretical results through a simulation study and on Malaria in Senegalese children dataset.

1. Introduction

The estimation of the probability density under censoring is a fundamental problem in statistics of considerable interest in many applied fields such as survival analysis, economics, engineering. Censoring is a condition in which the observation is only partially known. Let $T, T_1, \dots, T_n$ be independent, identically distributed random variables, and let f and F denote respectively the probability density and the distribution of T. In many situations, the full data $T_1, \dots, T_n$ are not available, let $C, C_1, \dots, C_n$ be the corresponding censoring random variables, and let G denote the distribution function of C. T and C are assumed to be independent. We consider here the two cases right and left censored data. The observed random variables are then X_i and δ_i where

$$X_i = \begin{cases} \min(T_i, C_i) & \text{and} \quad \delta_i = \mathbb{1}_{\{T_i \le C_i\}},\ 1 \le i \le n, \quad \text{right censoring} \\ \max(T_i, C_i) & \text{and} \quad \delta_i = \mathbb{1}_{\{T_i \ge C_i\}},\ 1 \le i \le n, \quad \text{left censoring,} \end{cases}$$

and we let $\pi_i = \mathbb{P}[\delta_i = 1 | T_i]$, this probability is called the propensity score (see [16], in the framework of missing data).

In this paper, we propose a data driven bandwidth selection for censored recursive kernel density estimators defined by stochastic approximation method. Recently, data driven bandwidth selection method for recursive kernel density estimators defined by stochastic approximation method have been investigated by [20] in the case of full data. [24] developed a data driven bandwidth selection

Received 2019-8-1; Accepted 2019-8-8; Communicated by the editors.
2010 *Mathematics Subject Classification.* Primary 62G07, 62L20; Secondary 65D10, 62N01.
Key words and phrases. Density estimation, stochastic approximation algorithm, smoothing, curve fitting, censored data.
* Corresponding author.

in the case of missing data, while [21] consider the recursive kernel distribution estimators. Recursive kernel regression estimators with a fixed design setting have been investigated by [22], while the semi-recursive kernel regression estimators have been studied by [23]. In this paper, we developed a specific data driven bandwidth selection method for the recursive kernel density estimators under censoring.

To construct a stochastic algorithm, which approximates the function f at a given point x, we define an algorithm of search of the zero of the function $g : y \to f(x) - y$. Following Robbins-Monro's scheme (see [15]), we set $\widehat{f}_0(x) \in \mathbb{R}$ and for all $n \geq 1$, $\widehat{f}_n(x) = \widehat{f}_{n-1}(x) + \gamma_n W_n(x)$, where (γ_n) is a nonrandom positive sequence tending to zero as n goes to infinity, called the stepsize. In order to define W_n at a point x we follow [13, 14], (see also [27] and [21]), and we introduce a kernel K (that is, a function satisfying $\int_{\mathbb{R}} K(x)dx = 1$), and a bandwidth (h_n) (that is, a sequence of positive real numbers that goes to zero), and we set $W_n(x) = h_n^{-1}\delta_n\widehat{\pi}_n^{-1}K\left(h_n^{-1}(x - X_n)\right) - f_{n-1}(x)$, where

$$\widehat{\pi}_i = \frac{\sum_{k=1}^n Q_k^{-1}\gamma_k h_k^{-1}\delta_k K\left(h_k^{-1}[X_i - X_k]\right)}{\sum_{k=1}^n Q_k^{-1}\gamma_k h_k^{-1} K\left(h_k^{-1}[X_i - X_k]\right)} \quad \text{with} \quad Q_n = \prod_{j=1}^n (1 - \gamma_j). \tag{1.1}$$

Then, our proposal in this paper to estimate recursively the function f at the point x is given by the following relation

$$\widehat{f}_n(x) = (1 - \gamma_n)\widehat{f}_{n-1}(x) + \gamma_n\delta_n\widehat{\pi}_n^{-1}h_n^{-1}K\left(h_n^{-1}[x - X_n]\right). \tag{1.2}$$

Throughout this paper, we suppose that $\widehat{f}_0(x) = 0$. Then, we can estimate f recursively at the point x by

$$\widehat{f}_n(x) \quad = \quad Q_n \sum_{k=1}^n Q_k^{-1}\gamma_k\delta_k\widehat{\pi}_k^{-1}h_k^{-1}K\left(\frac{x - X_k}{h_k}\right).$$

Moreover, one can check that by considering the quantity $\mathbb{E}\int_{\mathbb{R}}\left[\widehat{f}_n(x) - f(x)\right]^2 dx$ as criteria of selecting the optimal bandwidth, the optimal bandwidth depend on the choice of the stepsizes (γ_n); in particular under some conditions of regularity of f and G and using the classical stepsizes $(\gamma_n) = \left(n^{-1}\right)$, the bandwidth (h_n) is asymptotically

$$\left(\left(\frac{3}{10}\right)^{1/5}\left\{\int_{\mathbb{R}}\left(\Delta^{(2)}(x)\right)^2 dx\right\}^{-1/5}\left\{\frac{\int_{\mathbb{R}} K^2(z)\,dz}{\left(\int_{\mathbb{R}} z^2K(z)\,dz\right)^2}\right\}^{1/5}\widehat{\pi}_n^{1/5}n^{-1/5}\right),$$

where

$$\Delta(x) = \begin{cases} f(x)(1 - G(x)) & \text{right censoring} \\ f(x)G(x) & \text{left censoring.}\end{cases} \tag{1.3}$$

This considered bandwidth depend on the unknown quantity $\int_{\mathbb{R}}\left(\Delta^{(2)}(x)\right)^2 dx$, one can observe that is not easy to construct an asymptotic unbiased estimator of this quantity using just the observed data, rather than $\mathbb{E}\int_{\mathbb{R}}\left[\widehat{f}_n(x) - f(x)\right]^2 dx$, we propose to use as criteria of selecting the optimal bandwidth the quantity $\mathbb{E}\int_{\mathbb{R}}\left[\widehat{f}_n(x) - f(x)\right]^2 w(x)\,dx$, by considering the function $w(x) = \Delta^2(x)f^{-1}(x)$ as a weight function; which conducted to estimate the two quantities $\int_{\mathbb{R}}\Delta^2(x)\,dx$ and $\int_{\mathbb{R}}\left(\Delta^{(2)}(x)\right)^2\Delta(x)\,dx$, which can be estimated using just the observed data (see Eqs. (2.9) and (2.10)). Then, the first aim of this paper is to propose a second generation plug-in bandwidth selection, and the second aim is to give the conditions under which the recursive estimators $\widehat{f}_n$ can be better than the non-recursive kernel density estimator under censoring, and defined as

$$\widetilde{f}_n(x) = \frac{1}{nh_n}\sum_{k=1}^{n}\delta_k\widetilde{\pi}_k^{-1}K\left(\frac{x-X_k}{h_n}\right), \tag{1.4}$$

where

$$\widetilde{\pi}_i = \frac{\sum_{k=1}^{n}\delta_k K\left(h_n^{-1}\left[X_i-X_k\right]\right)}{\sum_{k=1}^{n}K\left(h_n^{-1}\left[X_i-X_k\right]\right)} \quad \text{(see [11, 28]).} \tag{1.5}$$

The numerical studies given in Section 3 are corroborating these theoretical results. The layout of the paper is as follows. In Section 2, we state our main results. Section 3 is devoted to our numerical studies, first by simulation (subsection 3.1) and second using a real dataset (subsection 3.3). We conclude the paper in Section 4. Appendix A gives the proof of our theoretical results.

2. Assumptions and Main Results

We define the following class of regularly varying sequences.

Definition 2.1. Let $\gamma\in\mathbb{R}$ and $(v_n)_{n\geq 1}$ be a nonrandom positive sequence. We say that $(v_n)\in\mathcal{GS}(\gamma)$ if

$$\lim_{n\to+\infty} n\left[1-\frac{v_{n-1}}{v_n}\right]=\gamma. \tag{2.1}$$

Condition (2.1) was introduced by [5] to define regularly varying sequences (see also [2]) and by [8] in the context of stochastic approximation algorithms. Noting that the acronym $\mathcal{GS}$ stand for (Galambos and Seneta). Typical sequences in $\mathcal{GS}(\gamma)$ are, for $b\in\mathbb{R}$, $n^{\gamma}(\log n)^{b}$, $n^{\gamma}(\log\log n)^{b}$, and so on.

In this section, we investigate asymptotic properties of the proposed estimators (1.2). The assumptions to which we shall refer are the following

- **:** (A1) $K:\mathbb{R}\to\mathbb{R}$ is a non-negative continuous, bounded function satisfying $\int_{\mathbb{R}}K(z)\,dz=1$, and, $\int_{\mathbb{R}}zK(z)\,dz=0$ and $\int_{\mathbb{R}}z^2K(z)\,dz<\infty$.
- **:** (A2) *i)* $(\gamma_n)\in\mathcal{GS}(-\alpha)$ with $\alpha\in(1/2,1]$.
  ii) $(h_n)\in\mathcal{GS}(-a)$ with $a\in(0,1)$.
  iii) $\lim_{n\to\infty}(n\gamma_n)\in(\min\{2a,(\alpha-a)/2\},\infty]$.
- **:** (A3) f is bounded, twice differentiable, and $f^{(2)}$ is bounded.

- The intuition behind the use of such bandwidth (h_n) belonging to $\mathcal{GS}(-a)$ is that the ratio h_{n-1}/h_n is equal to $1+a/n+o(1/n)$, then using such bandwidth and using the assumption $(A2)$ on the bandwidth and on the stepsize, Lemma A.2 ensures that the bias and the variance will depend only on h_n and not on $h_1,\ldots,h_n$, then the $MISE$ will depend also only on h_n, which will be helpful to deduce an optimal bandwidth.
- In order to help the readers to follow the main results obtained in this paper, we underline that under the assumption $(A2)$, we have $Q_n\sum_{k=1}^{n}Q_k^{-1}\gamma_k=1+o(1)$, $Q_n\sum_{k=1}^{n}Q_k^{-1}\gamma_k h_k^2=O\left(h_n^2\right)$ and $Q_n^2\sum_{k=1}^{n}Q_k^{-2}\gamma_k^2h_k^{-1}=O\left(\gamma_nh_h^{-1}\right)$.

- Assumption $(A2)\,(iii)$ on the limit of $(n\gamma_n)$ as n goes to infinity is usual in the framework of stochastic approximation algorithms. It implies in particular that the limit of $\left([n\gamma_n]^{-1}\right)$ is finite.

For simplicity, we introduce the following notations:

$$\begin{aligned}
\xi &= \lim_{n\to\infty}(n\gamma_n)^{-1}, && (2.2)\\
R(K) = \int_{\mathbb{R}} K^2(z)\,dz, \quad \mu_j(K) &= \int_{\mathbb{R}} z^j K(z)\,dz,\\
\Theta(K) &= R(K)^{4/5}\mu_2(K)^{2/5},\\
I_1 = \int_{\mathbb{R}} \Delta^2(x)\,dx, \quad I_2 &= \int_{\mathbb{R}} \left(\Delta^{(2)}(x)\right)^2 \Delta(x)\,dx,
\end{aligned}$$

where $\Delta^{(2)}$ represent the second derivative of Δ.

2.1. Results on the recursive estimators $\widehat{f}_n$. In this subsection, we explicit the choice of (h_n) through a second generation plug-in method, which is based on minimizing the Mean Weighted Integrated Squared Error ($MWISE$) of the proposed recursive estimator (1.2), in order to provide a comparison with the non-recursive estimator (1.4). Our first result is the following proposition, which gives the bias and the variance of $\widehat{f}_n$.

Proposition 2.2 (Bias and variance of $\widehat{f}_n$)**.** *Let Assumptions $(A1)-(A3)$ hold, and assume that $f^{(2)}$ is continuous at x.*

(1) If $0<a\leq\alpha/5$, then

$$\mathbb{E}\left[\widehat{f}_n(x)\right]-f(x)=h_n^2\frac{\pi_n^{-1}}{2(1-2a\xi)}\Delta^{(2)}(x)\,\mu_2(K)+o\left(h_n^2\right). \qquad (2.3)$$

If $\alpha/5<a<1$, then

$$\mathbb{E}\left[\widehat{f}_n(x)\right]-f(x)=o\left(\sqrt{\gamma_n h_n^{-1}}\right). \qquad (2.4)$$

(2) If $\alpha/5\leq a<1$, then

$$Var\left[\widehat{f}_n(x)\right]=\frac{\gamma_n}{h_n}\frac{\pi_n^{-1}}{(2-(\alpha-a)\xi)}f(x)\,R(K)+o\left(\frac{\gamma_n}{h_n}\right). \qquad (2.5)$$

If $0<a<\alpha/5$, then

$$Var\left[\widehat{f}_n(x)\right]=o\left(h_n^4\right). \qquad (2.6)$$

(3) If $\lim_{n\to\infty}(n\gamma_n)>\max\{2a,(\alpha-a)/2\}$, then (2.3) and (2.5) hold simultaneously.

The bias and the variance of the estimators $\widehat{f}_n$ defined by the stochastic approximation algorithm (1.2) then heavily depend on the choice of the stepsizes (γ_n).

Let us first state the following theorem, which gives the weak convergence rate of the estimators $\widehat{f}_n$ defined in (1.2).

Theorem 2.3 (Weak pointwise convergence rate)**.** *Let Assumptions $(A1)-(A3)$ hold, and assume that $f^{(2)}$ is continuous at x.*

(1) If there exists $c \geq 0$ *such that* $\gamma_n^{-1} h_n^5 \to c$*, then*

$$\sqrt{\gamma_n^{-1} h_n}\left(\widehat{f}_n(x) - f(x)\right)$$
$$\xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\sqrt{c}\pi_n^{-1}}{2(1-2a\xi)}\Delta^{(2)}(x)\mu_2(K), \frac{\pi_n^{-1}}{(2-(\alpha-a)\xi)} f(x) R(K)\right).$$

(2) If $\gamma_n^{-1} h_n^5 \to \infty$*, then*

$$\frac{1}{h_n^2}\left(\widehat{f}_n(x) - f(x)\right) \xrightarrow{\mathbb{P}} \frac{\pi_n^{-1}}{2(1-2a\xi)}\Delta^{(2)}(x)\mu_2(K),$$

where $\xrightarrow{\mathcal{D}}$ *denotes the convergence in distribution,* $\mathcal{N}$ *the Gaussian-distribution and* $\xrightarrow{\mathbb{P}}$ *the convergence in probability.*

The following theorem gives the strong pointwise convergence rate of the proposed recursive kernel density estimators under censoring.

Theorem 2.4 (Strong pointwise convergence rate)**.** *Let Assumptions* $(A1) - (A3)$ *hold, and assume that* $f^{(2)}$ *is continuous at* x.

(1) If there exists $c_1 \geq 0$ *such that* $\gamma_n^{-1} h_n^5 / \ln\left(\sum_{k=1}^n \gamma_k\right) \to c_1$*, then, with probability one, the sequence*

$$\left(\sqrt{\frac{\gamma_n^{-1} h_n}{2\ln\left(\sum_{k=1}^n \gamma_k\right)}}\left(\widehat{f}_n(x) - f(x)\right)\right)$$

is relatively compact and its limit set is the interval

$$\left[\frac{\pi_n^{-1}}{2(1-2a\xi)}\sqrt{\frac{c_1}{2}}\Delta^{(2)}(x)\mu_2(K) - \sqrt{\frac{\pi_n^{-1} f(x)}{(2-(\alpha-a)\xi)}R(K)},\right.$$
$$\left.\frac{\pi_n^{-1}}{2(1-2a\xi)}\sqrt{\frac{c_1}{2}}\Delta^{(2)}(x)\mu_2(K) + \sqrt{\frac{\pi_n^{-1} f(x)}{(2-(\alpha-a)\xi)}R(K)}\right].$$

(2) If $\gamma_n^{-1} h_n^5 / \ln\left(\sum_{k=1}^n \gamma_k\right) \to \infty$ *then with probability one,*

$$\lim_{n\to\infty} h_n^{-2}\left(\widehat{f}_n(x) - f(x)\right) = \frac{\pi_n^{-1}}{2(1-2a\xi)}\Delta^{(2)}(x).$$

As a criteria of selecting the optimal bandwidth, we use the $MWISE$, by taking the function $w(x) = \Delta^2(x) f^{-1}(x)$ as a weight function,

$$\begin{aligned} MWISE\left[\widehat{f}_n\right] &= \int_{\mathbb{R}}\left(\mathbb{E}\left(\widehat{f}_n(x)\right) - f(x)\right)^2 \Delta^2(x) f^{-1}(x)\,dx \\ &\quad + \int_{\mathbb{R}} Var\left(\widehat{f}_n(x)\right)\Delta^2(x) f^{-1}(x)\,dx. \end{aligned}$$

The following proposition gives the $MWISE$ of the proposed recursive estimators given in (1.2).

Proposition 2.5 ($MWISE$ of $\widehat{f}_n$)**.** *Let Assumptions* $(A1) - (A3)$ *hold, and assume that* $f^{(2)}$ *is continuous and integrable.*

(1) If $0 < a < \alpha/5$*, then*

$$MWISE\left[\widehat{f}_n\right] = \frac{1}{4} h_n^4 \frac{\pi_n^{-2}}{(1-2a\xi)^2} I_2 \mu_2^2(K) + o\left(h_n^4\right).$$

(2) *If* $a = \alpha/5$, *then*

$$\begin{aligned} MWISE\left[\widehat{f}_n\right] &= \frac{\gamma_n}{h_n}\frac{\pi_n^{-1}}{(2-(\alpha-a)\xi)}I_1R(K)+\frac{1}{4}h_n^4\frac{\pi_n^{-2}}{(1-2a\xi)^2}I_2\mu_2^2(K) \\ &\quad +o\left(h_n^4\right). \end{aligned}$$

(3) *If* $\alpha/5 < a < 1$, *then*

$$MWISE\left[\widehat{f}_n\right] = \frac{\gamma_n}{h_n}\frac{\pi_n^{-1}}{(2-(\alpha-a)\xi)}I_1R(K)+o\left(\frac{\gamma_n}{h_n}\right).$$

The following corollary ensures that the bandwidth which minimize the $MWISE$ of the proposed recursive estimator $\widehat{f}_n$ depend on the stepsize (γ_n) and on the propensity score π_n and then the corresponding $MWISE$ depend also on the stepsize (γ_n) and on the propensity score π_n.

Corollary 2.6. *Let Assumptions* $(A1)-(A3)$ *hold. To minimize the* $MWISE$ *of* $\widehat{f}_n$, *the stepsize* (γ_n) *must be chosen in* $\mathcal{GS}(-1)$, *the bandwidth* (h_n) *must equal*

$$\left(\left\{\frac{(1-2a\xi)^2}{(2-(1-a)\xi)}\frac{I_1}{I_2}\right\}^{1/5}\left\{\frac{R(K)}{\mu_2^2(K)}\right\}^{1/5}(\gamma_n\pi_n)^{1/5}\right).$$

Then, we have

$$\begin{aligned} MWISE\left[\widehat{f}_n\right] &= \frac{5}{4}(1-2a\xi)^{-2/5}(2-(1-a)\xi)^{-4/5}I_1^{4/5}I_2^{1/5}\Theta(K)\pi_n^{-6/5}\gamma_n^{4/5} \\ &\quad +o\left(\gamma_n^{4/5}\right). \end{aligned}$$

The following corollary shows that, for a special choice of the stepsize $(\gamma_n) = \left(\gamma_0 n^{-1}\right)$, which fulfilled that $\lim_{n\to\infty} n\gamma_n = \gamma_0$ and that $(\gamma_n) \in \mathcal{GS}(-1)$, the optimal value for the bandwidth (h_n) depend on γ_0 and then the corresponding $MWISE$ depend also on γ_0.

Corollary 2.7. *Let Assumptions* $(A1)-(A3)$ *hold, and suppose that* $(\gamma_n) = \left(\gamma_0 n^{-1}\right)$. *To minimize the* $MWISE$ *of* $\widehat{f}_n$, *the stepsize* (γ_n) *must be chosen in* $\mathcal{GS}(-1)$, $\lim_{n\to\infty} n\gamma_n = \gamma_0$, *the bandwidth* (h_n) *must equal*

$$\left(2^{-1/5}(\gamma_0-2/5)^{1/5}\left(\frac{I_1}{I_2}\right)^{1/5}\left\{\frac{R(K)}{\mu_2^2(K)}\right\}^{1/5}\pi_n^{1/5}n^{-1/5}\right),$$

and then the asymptotic $MWISE$ *of the proposed recursive estimator* $\widehat{f}_n$ *is equal to*

$$AMWISE\left[\widehat{f}_n\right] = \frac{5}{4}\frac{1}{2^{4/5}}\frac{\gamma_0^2}{(\gamma_0-2/5)^{6/5}}I_1^{4/5}I_2^{1/5}\Theta(K)\pi_n^{-6/5}n^{-4/5}.$$

Moreover, the minimum of $\gamma_0^2(\gamma_0-2/5)^{-6/5}$ *is reached at* $\gamma_0 = 1$, *then the bandwidth* (h_n) *must equal*

$$\left(\left(\frac{3}{10}\right)^{1/5}\left(\frac{I_1}{I_2}\right)^{1/5}\left\{\frac{R(K)}{\mu_2^2(K)}\right\}^{1/5}\pi_n^{1/5}n^{-1/5}\right), \tag{2.7}$$

and we then

$$AMWISE\left[\widehat{f}_n\right] = \frac{5}{4}\frac{1}{2^{4/5}}\left(\frac{5}{3}\right)^{6/5} I_1^{4/5} I_2^{1/5}\Theta\left(K\right)\pi_n^{-6/5}n^{-4/5}. \quad (2.8)$$

In order to investigate a data-driven bandwidth selection based on the optimal bandwidth (2.7), we must estimate the unknown quantities I_1 and I_2. For this purpose, we used the following kernel estimators:

$$\widehat{I}_1 = \frac{Q_n}{n}\sum_{i,k=1}^{n} Q_k^{-1}\beta_k b_k^{-1}\delta_k K_b\left(\frac{X_i - X_k}{b_k}\right), \quad (2.9)$$

$$\widehat{I}_2 = \frac{Q_n^2}{n}\sum_{\substack{i,j,k=1\\ j\neq k}}^{n} Q_j^{-1}Q_k^{-1}\beta_j'\beta_k' b_j'^{-3} b_k'^{-3}\delta_j\delta_k K_{b'}^{(2)}\left(\frac{X_i - X_j}{b_j'}\right) K_{b'}^{(2)}\left(\frac{X_i - X_k}{b_k'}\right), \quad (2.10)$$

where K_b and $K_{b'}$ are a kernels, b_n and b_n' are respectively the associated bandwidth (called pilot bandwidth) and β_n and β_n' are the two pilot stepsizes for the estimation of I_1 and I_2 respectively.

In practice, we take

$$b_n = n^{-\beta}\min\left\{\widehat{s}, \frac{Q_3 - Q_1}{1.349}\right\}, \quad \beta \in (0,1) \quad (2.11)$$

(see [18]) with $\widehat{s}$ the sample standard deviation, and Q_1, Q_3 denoting the first and third quartiles, respectively.

At this stage, we need to give the optimal choice of (b_n), (b_n'), (β_n) and (β_n'). In order to achieve this task, we followed the same steps as the work of [20] in the context of recursive kernel density estimation with no censoring data, and we showed that in order to minimize the $AMISE$ of $\widehat{I}_1$ the pilot bandwidth (b_n) must belong to $\mathcal{GS}\left(-2/5\right)$ and the pilot stepsize (β_n) should be equal to $\left(1.36n^{-1}\right)$, and in order to minimize the $AMISE$ of $\widehat{I}_2$ the pilot bandwidth (b_n') must belong to $\mathcal{GS}\left(-3/14\right)$ and the pilot stepsize (β_n') should be equal to $\left(1.48n^{-1}\right)$.

Then, the conducted data-driven bandwidth selector (h_n) in this work using the proposed recursive estimators given in (1.2) with the chosen stepsizes $(\gamma_n) = \left(n^{-1}\right)$ is equal to

$$\left(\left(\frac{3}{10}\right)^{1/5}\left(\frac{\widehat{I}_1}{\widehat{I}_2}\right)^{1/5}\left\{\frac{R\left(K\right)}{\mu_2^2\left(K\right)}\right\}^{1/5}\widehat{\pi}_n^{1/5}n^{-1/5}\right), \quad (2.12)$$

and the associated plug-in $AMWISE$ is equal to

$$\widehat{AMWISE}\left[\widehat{f}_n\right] = \frac{5}{4}\frac{1}{2^{4/5}}\left(\frac{5}{3}\right)^{6/5}\widehat{I}_1^{4/5}\widehat{I}_2^{1/5}\Theta\left(K\right)\widehat{\pi}_n^{-6/5}n^{-4/5}.$$

2.2. Results on the non-recursive estimator $\widetilde{f}_n$. Let us claimed the following Lemma which gives the bias and variance of the non-recursive kernel density estimator under censoring $\widetilde{f}_n$, the proof follows immediately from the proof the Proposition 2.2.

Lemma 2.8 (Bias and variance of $\widetilde{f}_n$)**.** *Let Assumptions* $(A1)$, $(A2)\,ii)$ *and* $(A3)$ *hold, and assume that* $f^{(2)}$ *is continuous at* x.

$$\mathbb{E}\left[\widetilde{f}_n(x)\right] - f(x) = \frac{h_n^2}{2}\pi_n^{-1}\Delta^{(2)}(x)\,\mu_2(K) + o\left(h_n^2\right),$$

and

$$Var\left[\widetilde{f}_n(x)\right] = \frac{\pi_n^{-1}}{nh_n} f(x)\,R(K) + o\left(\frac{1}{nh_n}\right),$$

It follows from Lemma 2.8, that

$$AMWISE\left[\widetilde{f}_n\right] = \frac{\pi_n^{-1}}{nh_n} I_1 R(K) + \frac{h_n^4}{4}\pi_n^{-2} I_2 h_n^4 \mu_2^2(K).$$

Then, to minimize the $AMWISE$ of $\widetilde{f}_n$, the bandwidth (h_n) must equal to

$$\left(\left(\frac{I_1}{I_2}\right)^{1/5}\left\{\frac{R(K)}{\mu_2^2(K)}\right\}^{1/5}\pi_n^{1/5} n^{-1/5}\right), \tag{2.13}$$

and we have

$$AMWISE\left[\widetilde{f}_n\right] = \frac{5}{4} I_1^{4/5} I_2^{1/5}\Theta(K)\,\pi_n^{-6/5} n^{-4/5}. \tag{2.14}$$

Remark 2.9. It follows from (2.8) and (2.14), that the $AMWISE$ of the proposed recursive estimators $\widehat{f}_n$ with the choice of the stepsizes $(\gamma_n) = \left(n^{-1}\right)$ is 1.06 larger than the $AMWISE$ of the non-recursive estimator $\widetilde{f}_n$.

Moreover, in order to estimate the optimal bandwidth (2.13), we must estimate I_1 and I_2. For this purpose, we use the following two kernel estimators :

$$\widetilde{I}_1 = \frac{1}{n(n-1)\,b_n}\sum_{\substack{i,j=1\\ i\neq j}}^{n} \delta_j K_b\left(\frac{X_i - X_j}{b_n}\right), \tag{2.15}$$

$$\widetilde{I}_2 = \frac{1}{n^3 b_n'^6}\sum_{\substack{i,j,k=1\\ j\neq k}}^{n} \delta_j\delta_k K_{b'}^{(2)}\left(\frac{X_i - X_j}{b_n'}\right) K_{b'}^{(2)}\left(\frac{X_i - X_k}{b_n'}\right), \tag{2.16}$$

where K_b and $K_{b'}$ are a kernels, b_n and b_n' are respectively the associated bandwidth given in (2.11).

At this stage, we need to give the optimal choice of (b_n) and (b_n'). For thus purpose, we used as a criteria the $AMISE$ of $\widetilde{I}_1$ respectively of $\widetilde{I}_2$, we showed that, the pilot bandwidth (b_n) respectively (b_n') must belong to $\mathcal{GS}(-2/5)$, respectively to $\mathcal{GS}(-3/14)$.

Then, the data-driven bandwidth selector (h_n) using the non-recursive estimator (1.4), is given by

$$\left(\left(\frac{\widetilde{I}_1}{\widetilde{I}_2}\right)^{1/5}\left\{\frac{R(K)}{\mu_2^2(K)}\right\}^{1/5}\widetilde{\pi}_n^{1/5} n^{-1/5}\right), \tag{2.17}$$

and the plug-in of the $AMWISE$ of the non-recursive estimator (1.4), is given by

$$\widetilde{AMWISE}\left[\widetilde{f}_n\right] = \frac{5}{4}\widetilde{I}_1^{4/5}\widetilde{I}_2^{1/5}\Theta(K)\,\widetilde{\pi}_n^{-6/5} n^{-4/5}.$$

Moreover, a hint of the proof of the following corollary are given in the appendix.

Corollary 2.10. *Let the assumptions* $(A1)-(A3)$ *hold, and the bandwidth* (h_n) *equal to (2.12) and the stepsize* $(\gamma_n) = (n^{-1})$ *when we apply the estimators* $\widehat{f}_n$ *and the bandwidth* (h_n) *equal to (2.17) when we apply the estimator* $\widetilde{f}_n$*. We have*

$$\frac{\mathbb{E}\left[\widehat{MWISE}\left(\widehat{f}_n\right)\right]}{\mathbb{E}\left[\widetilde{MWISE}\left(\widetilde{f}_n\right)\right]} < 1 \quad \textit{for small sample setting.} \tag{2.18}$$

Then, for small sample size, the expected value of $\widehat{MWISE}$ *of the proposed recursive estimators defined by (1.2) where the bandwidth is replaced by its data-based version is smaller than the expected value of* $\widetilde{MWISE}$ *of the non-recursive estimator defined by (1.4) where the bandwidth is replaced by its data-based version.*

3. Numerical Studies

The aim of this section is to compare the performance of the proposed recursive kernel density estimators under censoring defined in (1.2) with that of the non-recursive estimator defined in (1.4). The comparison is done first through a simulation study and then through a real data set.

: When applying $\widehat{f}_n$ one need to choose four quantities:
  (1) The function K, we choose the Normal kernel.
  (2) The stepsizes (γ_n) equal to (n^{-1}).
  (3) The propensity score (π_n) is chosen to be equal to (1.1).
  (4) The bandwidth (h_n) is chosen to be equal to (2.12).
    (a) To estimate I_1, we use (2.9); The pilot bandwidth is chosen to be equal to (2.11) with the choice of $\beta = 2/5$ and the pilot stepsize equal to $(1.36n^{-1})$.
    (b) To estimate I_2, we use (2.10); The pilot bandwidth is chosen to be equal to (2.11) with the choice of $\beta = 3/14$ and the pilot stepsize equal to $(1.48n^{-1})$.

: When applying $\widetilde{f}_n$ one need to choose three quantities:
  (1) The function K, as in the recursive framework, we use the Normal kernel.
  (2) The propensity score (π_n) is chosen to be equal to (1.5).
  (3) The bandwidth (h_n) is chosen to be equal to (2.17).
    (a) To estimate I_1, we use (2.15); The pilot bandwidth is chosen to be equal to (2.11) with the choice of $\beta = 2/5$.
    (b) To estimate I_2, we use (2.16); The pilot bandwidth is chosen to be equal to (2.11) with the choice of $\beta = 3/14$.

3.1. Simulation experiments. In our simulation study, we consider three sample size, $n = 100$, $n = 200$, $n = 500$ and the following five densities functions f : 1- the standard normal : $X \sim \mathcal{N}(0,1)$, 2- the normal mixture distribution : $X \sim 1/2\mathcal{N}(2,1)+1/2\mathcal{N}(-3,1)$, 3- the weibull distribution with shape parameter 2 and scale parameter 1: $X \sim \mathcal{W}eibul(2,1)$, 4- the log normal distribution: $X \sim \log\mathcal{N}(0,1)$, 5- the chi squared distribution with 12 degree of freedom: $X \sim$

$\chi^2(12)$. Moreover, we simulated different censoring levels, in each considered case, we fixed the level after ordering the data. For each density and sample size n, we approximate the average ISE of the estimators using $N = 500$ trials of sample size n; $\overline{ISE} = \frac{1}{N}\sum_{k=1}^{N} ISE\left[\widehat{g}^{[k]}\right]$, where $\widehat{g}^{[k]}(.)$ is the estimator computed from the kth sample and $ISE\left[\widehat{g}^{[k]}\right] = \int_{\mathbb{R}}\left\{\widehat{g}^{[k]}(x) - f(x)\right\}dx$, and we approximate the average correlation; $\overline{Cor} = \frac{1}{N}\sum_{k=1}^{N} Cor\left(\widehat{g}^{[k]}, f(x)\right)$.

3.2. Computational cost. In order to give some comparative elements with the non-recursive censored estimator (1.4), including computational costs. We consider a 500 samples of size $n_1 = \lfloor n/2 \rfloor$ (the lower integer part of $n/2$) generated from respectively the five considered distributions, moreover, we suppose that we receive an additional 500 samples of size $n - n_1$ generated also from the same five considered densities.

This property can be generalized, one can check that it follows from (1.2) that for all $n_1 \in [0, n-1]$,

$$\begin{aligned}
\widehat{f}_n(x) &= \prod_{j=n_1+1}^{n}(1-\gamma_j)\widehat{f}_{n_1}(x) \\
&\quad + \sum_{k=n_1}^{n-1}\left[\prod_{j=k+1}^{n}(1-\gamma_j)\right]\frac{\gamma_k}{h_k}\delta_k\widehat{\pi}_k^{-1}K\left(\frac{x-X_k}{h_k}\right) \\
&\quad + \frac{\gamma_n}{h_n}\delta_n\widehat{\pi}_n^{-1}K\left(\frac{x-X_n}{h_n}\right) \\
&= \alpha_1\widehat{f}_{n_1}(x) + \sum_{k=n_1}^{n-1}\beta_k\frac{\gamma_k}{h_k}\delta_k\widehat{\pi}_k^{-1}K\left(\frac{x-X_k}{h_k}\right) \\
&\quad + \frac{\gamma_n}{h_n}\delta_n\widehat{\pi}_n^{-1}K\left(\frac{x-X_n}{h_n}\right),
\end{aligned}$$

where $\alpha_1 = \prod_{j=n_1+1}^{n}(1-\gamma_j)$ and $\beta_k = \prod_{j=k+1}^{n}(1-\gamma_j)$. It is clear, that the proposed estimators can be viewed as a linear combination of two estimators, which improve considerably the computational cost.

Performing the two methods, we report the total `CPU` time values for each considered density and in all cases given in Tables 1, 2, 3, 4 and 5, the `CPU` time is given in seconds.

From Figures 1 and 2, Tables 1, 2, 3, 4 and 5, we conclude that

- The proposed recursive kernel density estimator under censoring (1.2), with the stepsize $(\gamma_n) = \left(n^{-1}\right)$ is closer to the true density function as compared with the non-recursive estimator (1.4).
- In all the considered densities, the proposed estimator performed better than the non-recursive kernel density (1.4) for a completed data (no censoring).
- In all the considered cases of tables 1, 3 and 4, the average ISE of the proposed recursive censored kernel density estimators is smaller than to the non-recursive kernel density estimator under censoring (1.4), even when the censoring levels is equal to 30% and the sample size equal to 500.

- In all the considered cases of table 2 and 5, the average ISE of the proposed recursive censored kernel density estimator using the stepsize $(\gamma_n) = (n^{-1})$ is quite similar to the non-recursive kernel density estimator under censoring (1.4).
- The estimators get closer to the true density function as sample size increase and the censoring level decrease.
- The two estimators ((1.2) and (1.4)) can deal effectively with both right and left censored data and under uncensored data.
- The `CPU` time are always faster using the proposed recursive estimator and the reduction of `CPU` time goes from a minimum of 32% to a maximum of 59% compared to the non-recursive estimator.

3.3. Real dataset. We considered a dataset of 176 families in Senegal, totalizing 505 children between 2 and 19 years old, living in two villages of Niakhar (Diohine and Toucar). The number of observations was 6986. We measured `Plasmodium falciparum` Parasite Load (`PL`) from thick blood smears obtained by finger-prick during two different seasons and regularly over a three-year observation period (2001-2003), the number of measurements per child ranged from 1 to 15, for more details see ([7]), this data was used also in [26] in a parametric context.

We had the following variables: 1- `Family identification` : A factor with 176 levels; 2- `Child identification` : A factor with 505 levels; 3- `PL` : Parasite Load; 4- `infection` : A factor with two levels (infected: 1 or not infected: 0); 5- `year` : A factor with three levels (0 for 2001, 1 for 2002 and 2 for 2003); 6- `number of measurements per child` : A factor with 15 levels; 7- `age` : Age of the child in years between 2 and 19; 8- `season` : A factor with two levels (July-October and October-March); 9- `village` : A factor with two levels (Diohine and Toucar).

Assuming that all these measurement are corrects, we simulated different censoring levels, in each considered case, we fixed the level after ordering the data.

The density are then compared to those obtained with the full data. Even when 10%, 20% or even 30% of the original measurements are censored, the produced density remain very accurate thus demonstrating the effectiveness of our approach.

Figure 3 illustrate the outcomes of the two estimators (1.2) and (1.4) compared to the non-recursive kernel estimator introduced by [17] (see also [12]) and the recursive estimator proposed in [10] using the stepsize $(\gamma_n) = (n^{-1})$ (see also [19] and [20]).

We refer by `N-rec` to the non-recursive estimator and by `Rec` to the recursive estimator.

4. Conclusion

This paper propose an automatic bandwidth selection of the recursive kernel density estimators under censoring defined by the stochastic approximation algorithm (1.2). The proposed estimators are consistent and asymptotically follows normal distribution. Our proposal estimators are compared to the non-recursive censored kernel density estimator (1.4). We showed that, using some data based

$X \sim \mathcal{N}(0,1)$

	$n=100$			$n=200$			$n=500$		
0%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$2.5e^{-4}$	0.998	113	$1.5e^{-4}$	0.999	484	$6.0e^{-5}$	0.999	2062
Rec	$2.3e^{-4}$	0.999	57	$1.4e^{-4}$	0.999	293	$5.9e^{-5}$	0.999	1320
10%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$2.4e^{-3}$	0.979	131	$2.2e^{-3}$	0.978	477	$1.9e^{-3}$	0.977	2247
Rec	$2.4e^{-3}$	0.980	58	$2.1e^{-3}$	0.979	258	$1.9e^{-3}$	0.978	1308
20%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$8.7e^{-3}$	0.899	132	$7.8e^{-3}$	0.906	483	$6.8e^{-3}$	0.911	2046
Rec	$8.5e^{-3}$	0.906	56	$7.7e^{-3}$	0.909	228	$6.7e^{-3}$	0.914	1228
30%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$1.7e^{-2}$	0.727	130	$1.6e^{-2}$	0.748	484	$1.4e^{-2}$	0.763	2060
Rec	$1.7e^{-2}$	0.735	58	$1.6e^{-2}$	0.754	286	$1.4e^{-2}$	0.770	1034

TABLE 1. Average $ISEs$ and Correlations (approximated using $N=500$ trials) and total CPU time in seconds of the non-recursive estimator $\widetilde{f}_n$ and the proposed recursive estimator $\widehat{f}_n$ with the choice of the stepsize $(\gamma_n)=\left(n^{-1}\right)$. Here we consider the normal distribution $X \sim \mathcal{N}(0,1)$ with the censoring level equal respectively to 0% (in the first block), equal to 10% (in the second block), equal to 20% (in the third block) and equal to 30% (in the last block), we consider three sample sizes $n=100$, $n=200$ and $n=500$.

selected bandwidth and some particularly stepsizes, the proposed recursive estimator can be better than the non-recursive one in terms of estimation error. The two estimators ((1.2) and (1.4)) can deal effectively with both right and left censored and uncensored data. The simulation study confirms the nice features of our proposed recursive estimators and satisfactory improvement in the CPU time in comparison to the non-recursive estimator.

In conclusion, the proposed method allowed us to obtain better results compared to the non-recursive censored kernel density estimator in terms of estimation error and much better in terms of computational costs. Moreover, we plan to consider the following estimator computed with Kaplan-Meier (KM) weights to estimate a density of probability in the presence of censored data,

$$\overline{f}_n(x) \quad = \quad \Pi_n \sum_{k=1}^{n} \Pi_k^{-1} \gamma_k h_k^{-1} \delta_k G_n^{-1}(X_k) K\left(\frac{x-X_k}{h_k}\right). \tag{4.1}$$

where $G_n(.)$ is given by

$$G_n(t) = \begin{cases} \prod_{i=1}^{n}\left(1-\frac{1-\delta_{(i)}}{n-i+1}\right)^{\mathbb{1}\left\{X_{(i)} \leq t\right\}} & \text{if} \quad t < X_{(n)} \\ 0 & \text{otherwise} \end{cases},$$

which is known to be uniformly convergent to G and $X_{(1)} < X_{(2)} < \cdot < X_{(n)}$ are the order statistics of $(X_{(i)})_{1 \leq i \leq n}$ and $\delta_{(i)}$ is the concomitant of $X_{(i)}$, and then to compare the estimator (4.1) to the kernel density estimator computed with Kaplan-Meier, see [1], [4] and [3], and then to compare the two estimators to

$X \sim \frac{1}{2}\mathcal{N}(2,1) + \frac{1}{2}\mathcal{N}(-3,1)$

	$n=100$			$n=200$			$n=500$		
0%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$1.1e^{-4}$	0.998	136	$6.8e^{-5}$	0.999	499	$3.3e^{-5}$	0.999	2142
Rec	$1.0e^{-4}$	0.998	70	$6.5e^{-5}$	0.999	266	$3.2e^{-5}$	0.999	1118
10%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$5.4e^{-4}$	0.955	133	$5.1e^{-4}$	0.959	481	$4.9e^{-4}$	0.961	2032
Rec	$5.5e^{-4}$	0.955	69	$5.1e^{-4}$	0.959	261	$4.9e^{-4}$	0.961	1070
20%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$2.0e^{-3}$	0.836	138	$1.9e^{-3}$	0.838	493	$1.8e^{-3}$	0.845	2018
Rec	$2.0e^{-3}$	0.823	76	$1.9e^{-3}$	0.837	259	$1.8e^{-3}$	0.845	1069
30%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$4.1e^{-3}$	0.645	134	$3.8e^{-3}$	0.669	491	$3.7e^{-3}$	0.684	2041
Rec	$4.1e^{-3}$	0.639	71	$3.8e^{-3}$	0.665	251	$3.7e^{-3}$	0.683	1069

Table 2. Average $ISEs$ and Correlations (approximated using $N = 500$ trials) and total CPU time in seconds of the non-recursive estimator $\widetilde{f}_n$ and the proposed recursive estimator $\widehat{f}_n$ with the choice of the stepsize $(\gamma_n) = (n^{-1})$. Here we consider the normal mixture distribution $X \sim \frac{1}{2}\mathcal{N}(2,1) + \frac{1}{2}\mathcal{N}(-3,1)$, with the censoring level equal respectively to 0% (in the first block), equal to 10% (in the second block), equal to 20% (in the third block) and equal to 30% (in the last block), we consider three sample sizes $n = 100$, $n = 200$ and $n = 500$.

those conducted in this work. We plan also to extend this work by considering Bernstein polynomials rather than kernels and to propose an adaptation of the estimators developed in [6] and [25] in the case of censored data.

Appendix A. Proofs

First, we approximate the estimators $\widehat{f}_n$ by the unobservable estimators f_n recursively defined by

$$f_n(x) = (1-\gamma_n) f_{n-1}(x) + \gamma_n \delta_n \pi_n^{-1} h_n^{-1} K\left(h_n^{-1}[x - X_n]\right). \tag{A.1}$$

Remark A.1. The consistency results of $\widehat{\pi}_n$ can be obtained from the generalized recursive version of Nadaraya-Watson's estimator proposed in ([23])).

Throughout this section we use the following notation:

$$\begin{aligned} Q_n &= \prod_{j=1}^{n}(1-\gamma_j), \\ Z_n(x) &= h_n^{-1}\delta_n\pi_n^{-1}K\left(\frac{x-X_n}{h_n}\right). \end{aligned} \tag{A.2}$$

Before giving the outlines of the proofs, we state the following technical lemma, which is proved in [10], and which will be used throughout the demonstrations.

$X \sim \mathcal{W}eibull\,(2,1)$

	$n=100$			$n=200$			$n=500$		
0%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$4.7e^{-4}$	0.999	133	$2.3e^{-4}$	0.999	488	$1.1e^{-4}$	0.999	2028
Rec	$4.6e^{-4}$	0.999	87	$2.2e^{-4}$	0.999	254	$1.1e^{-4}$	0.999	1085
10%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$4.9e^{-3}$	0.994	138	$4.9e^{-3}$	0.992	494	$4.7e^{-3}$	0.990	2056
Rec	$4.9e^{-3}$	0.994	69	$4.8e^{-3}$	0.992	256	$4.6e^{-3}$	0.990	1048
20%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$1.8e^{-2}$	0.970	137	$1.7e^{-2}$	0.964	487	$1.8e^{-2}$	0.957	2044
Rec	$1.8e^{-2}$	0.970	80	$1.7e^{-2}$	0.964	278	$1.7e^{-2}$	0.958	1068
30%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$4.0e^{-2}$	0.920	134	$3.8e^{-2}$	0.910	491	$3.7e^{-2}$	0.900	2051
Rec	$3.6e^{-2}$	0.921	92	$3.7e^{-2}$	0.913	256	$3.6e^{-2}$	0.902	1080

Table 3. Average $ISEs$ and Correlations (approximated using $N = 500$ trials) and total CPU time in seconds of the non-recursive estimator $\widetilde{f}_n$ and the proposed recursive estimator $\widehat{f}_n$ with the choice of the stepsize $(\gamma_n) = (n^{-1})$. Here we consider the weibull distribution with shape parameter 2 and scale parameter 1, $X \sim \mathcal{W}eibull\,(2,1)$, with the censoring level equal respectively to 0% (in the first block), equal to 10% (in the second block), equal to 20% (in the third block) and equal to 30% (in the last block), we consider three sample sizes $n = 100$, $n = 200$ and $n = 500$.

Lemma A.2. *Let $(v_n) \in \mathcal{GS}(v^*)$, $(\eta_n) \in \mathcal{GS}(-\eta)$, and $m > 0$ such that $m - v^*\xi > 0$ where ξ is defined in* (2.2)*. We have*

$$\lim_{n\to+\infty} v_n Q_n^m \sum_{k=1}^n Q_k^{-m}\gamma_k v_k^{-1} = (m - v^*\xi)^{-1}.$$

Moreover, for all positive sequence (α_n) such that $\lim_{n\to+\infty}\alpha_n = 0$, and all $C \in \mathbb{R}$,

$$\lim_{n\to+\infty} v_n Q_n^m \left[\sum_{k=1}^n Q_k^{-m}\eta_k v_k^{-1}\alpha_k + C\right] = 0.$$

Lemma A.2 is widely applied throughout the proofs. Let us underline that it is its application, which requires Assumption $(A2)(iii)$ on the limit of $(n\gamma_n)$ as n goes to infinity.

Our proofs are organized as follows. Propositions 2.2 and 2.5 in Sections A.1 and A.2 respectively, Theorem 2.3 in Section A.3.

A.1. Proof of Proposition 2.2. In view of (A.1) and (A.2), we have

$$\begin{aligned} f_n(x) - f(x) &= Q_n \sum_{k=1}^n Q_k^{-1}\gamma_k \left(Z_k(x) - f(x)\right) \\ &\quad + Q_n\left(f_0(x) - f(x)\right). \end{aligned} \tag{A.3}$$

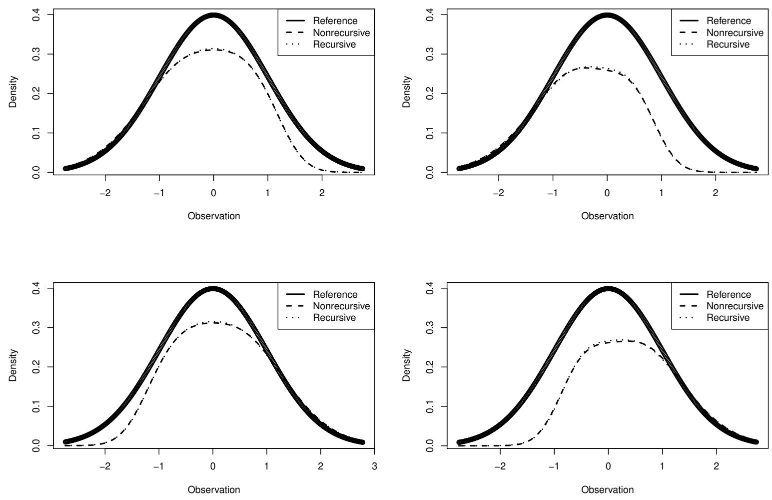

FIGURE 1. Qualitative comparison between the non-recursive kernel density estimator (1.4) (approximated using $N = 500$ trials of size $n = 200$ and given by the dashed line) and the recursive kernel density estimator (1.2) (approximated using $N = 500$ trials of size $n = 200$ and given by the dotted line), with the right censoring level equal respectively to 15% (in the top left panel), equal to 30% (in the top right panel), and with the left censoring level equal to 15% (in the down left panel) and equal to 30% (in the down-right panel) for the normal distribution $\mathcal{N}(0,1)$.

It follows that

$$\begin{aligned}\mathbb{E}\left(f_n\left(x\right)\right)-f\left(x\right) &= Q_n\sum_{k=1}^{n}Q_k^{-1}\gamma_k\left(\mathbb{E}\left(Z_k\left(x\right)\right)-f\left(x\right)\right)\\ &\quad +Q_n\left(f_0\left(x\right)-f\left(x\right)\right).\end{aligned}$$

First in the case of right censoring and since $X_k = \min\left(T_k, C_k\right)$, we have

$$\begin{aligned}\mathbb{E}\left[Z_k^p\left(x\right)\right] &= h_k^{-p}\pi_k^{-p}\mathbb{E}\left[\mathbb{1}_{\{T_k\leq C_k\}}K^p\left(\frac{x-X_k}{h_k}\right)\right]\\ &= h_k^{-p}\pi_k^{-p}\mathbb{E}\left[\mathbb{1}_{\{T_k\leq C_k\}}K^p\left(\frac{x-T_k}{h_k}\right)\right]\\ &= h_k^{-p}\pi_k^{-p}\int_{\mathbb{R}}\int_{\mathbb{R}}\mathbb{1}_{\{t\leq c\}}K^p\left(\frac{x-t}{h_k}\right)f\left(t\right)G'\left(c\right)dtdc\\ &= h_k^{-p}\pi_k^{-p}\int_{\mathbb{R}}\left(\int_{\mathbb{R}}\mathbb{1}_{\{t\leq c\}}G'\left(c\right)dc\right)K^p\left(\frac{x-t}{h_k}\right)f\left(t\right)dt\\ &= h_k^{-p}\pi_k^{-p}\int_{\mathbb{R}}\left(1-G\left(t\right)\right)K^p\left(\frac{x-t}{h_k}\right)f\left(t\right)dt. \qquad \text{(A.4)}\end{aligned}$$

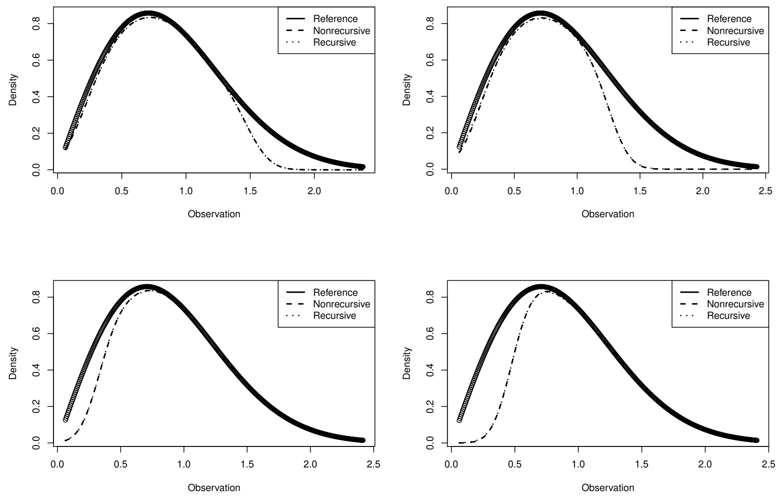

FIGURE 2. Qualitative comparison between the non-recursive kernel density estimator (1.4) (approximated using $N = 500$ trials of size $n = 200$ and given by the dashed line) and the recursive kernel density estimator (1.2) (approximated using $N = 500$ trials of size $n = 200$ and given by the dotted line), with the right censoring level equal respectively to 15% (in the top left panel), equal to 30% (in the top right panel), and with the left censoring level equal to 15% (in the down left panel) and equal to 30% (in the down-right panel) for the Weibull distribution with shape parameter 2 and scale parameter 1, $X \sim \mathcal{W}eibull\,(2,1)$.

Now, in the case of left censoring and since $X_k = \max(T_k, C_k)$, we have

$$\begin{aligned}
\mathbb{E}\left[Z_k^p(x)\right] &= h_k^{-p}\pi_k^{-p}\mathbb{E}\left[\mathbb{1}_{\{T_k\geq C_k\}}K^p\left(\frac{x-X_k}{h_k}\right)\right]\\
&= h_k^{-p}\pi_k^{-p}\mathbb{E}\left[\mathbb{1}_{\{T_k\geq C_k\}}K^p\left(\frac{x-T_k}{h_k}\right)\right]\\
&= h_k^{-p}\pi_k^{-p}\int_{\mathbb{R}}\int_{\mathbb{R}}\mathbb{1}_{\{t\geq c\}}K^p\left(\frac{x-t}{h_k}\right)f(t)\,G'(c)\,dtdc\\
&= h_k^{-p}\pi_k^{-p}\int_{\mathbb{R}}\left(\int_{\mathbb{R}}\mathbb{1}_{\{t\geq c\}}G'(c)\,dc\right)K^p\left(\frac{x-t}{h_k}\right)f(t)\,dt\\
&= h_k^{-p}\pi_k^{-p}\int_{\mathbb{R}}G(t)\,K^p\left(\frac{x-t}{h_k}\right)f(t)\,dt. \qquad (A.5)
\end{aligned}$$

Then, in follows from (1.3), (A.4) and (A.5), that

$$\mathbb{E}\left[Z_k^p(x)\right] = h_k^{-p+1}\pi_k^{-p}\int_{\mathbb{R}}\Delta\left(x-zh_k\right)K^p(z)\,dz. \qquad (A.6)$$

$X \sim \log\mathcal{N}(0,1)$

	$n=100$			$n=200$			$n=500$		
0%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$2.3e^{-3}$	0.9623	139	$1.5e^{-3}$	0.968	498	$8.7e^{-4}$	0.977	2069
Rec	$2.1e^{-3}$	0.9659	69	$1.4e^{-3}$	0.971	292	$8.1e^{-4}$	0.979	1045
10%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$3.1e^{-3}$	0.951	153	$2.2e^{-3}$	0.958	479	$1.3e^{-3}$	0.967	2066
Rec	$2.8e^{-3}$	0.956	72	$2.0e^{-3}$	0.963	266	$1.2e^{-3}$	0.970	1073
20%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$4.3e^{-3}$	0.946	164	$3.0e^{-3}$	0.949	489	$1.8e^{-3}$	0.955	2081
Rec	$4.0e^{-3}$	0.953	86	$2.8e^{-3}$	0.955	253	$1.7e^{-3}$	0.960	1053
30%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$5.6e^{-3}$	0.950	168	$4.1e^{-3}$	0.948	522	$2.7e^{-3}$	0.946	2134
Rec	$5.4e^{-3}$	0.957	69	$3.9e^{-3}$	0.955	287	$2.6e^{-3}$	0.952	1070

TABLE 4. Average $ISEs$ and Correlations (approximated using $N = 500$ trials) and total CPU time in seconds of the non-recursive estimator $\widetilde{f}_n$ and the proposed recursive estimator $\widehat{f}_n$ with the choice of the stepsize $(\gamma_n) = \left(n^{-1}\right)$. Here we consider the log normal distribution $X \sim \log\mathcal{N}(0,1)$, with the censoring level equal respectively to 0% (in the first block), equal to 10% (in the second block), equal to 20% (in the third block) and equal to 30% (in the last block), we consider three sample sizes $n = 100$, $n = 200$ and $n = 500$.

Then, it follows from (A.6), for $p = 1$, that

$$\begin{aligned}\mathbb{E}\left[Z_k(x)\right] - f(x) &= \pi_k^{-1}\int_{\mathbb{R}} K(z)\left[\Delta(x - zh_k) - \Delta(x)\right]dz \\ &= \pi_k^{-1}\frac{h_k^2}{2}\Delta^{(2)}(x)\,\mu_2(K) + \eta_k(x),\end{aligned} \tag{A.7}$$

with

$$\eta_k(x) = \pi_k^{-1}\int_{\mathbb{R}} K(z)\left[\Delta(x - zh_k) - \Delta(x) - \frac{1}{2}z^2h_k^2\Delta^{(2)}(x)\right]dz,$$

and, since Δ is bounded and continuous at x, we have $\lim_{k\to\infty}\eta_k(x) = 0$. In the case $a \leq \alpha/5$, we have $\lim_{n\to\infty}(n\gamma_n) > 2a$; the application of Lemma A.2 then gives

$$\begin{aligned}\mathbb{E}\left[f_n(x)\right] - f(x) &= \frac{1}{2}\pi_n^{-1}\Delta^{(2)}(x)\int_{\mathbb{R}} z^2K(z)\,dz\,Q_n\sum_{k=1}^{n}Q_k^{-1}\gamma_k h_k^2\left[1 + o(1)\right] \\ &\quad + Q_n\left(f_0(x) - f(x)\right) \\ &= \frac{1}{2(1 - 2a\xi)}\pi_n^{-1}\Delta^{(2)}(x)\,\mu_2(K)\left[h_n^2 + o(1)\right],\end{aligned}$$

$X \sim \chi^2(12)$

	$n = 100$			$n = 200$			$n = 500$		
0%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$1.0e^{-5}$	0.997	130	$7.1e^{-6}$	0.998	493	$3.33e^{-6}$	0.999	2066
Rec	$9.5e^{-6}$	0.998	69	$6.7e^{-6}$	0.998	258	$3.25e^{-6}$	0.999	1095
10%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$3.4e^{-5}$	0.992	136	$3.3e^{-5}$	0.991	490	$3.1e^{-5}$	0.990	2149
Rec	$3.4e^{-5}$	0.992	70	$3.3e^{-5}$	0.991	256	$3.1e^{-5}$	0.990	1064
20%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$1.2e^{-4}$	0.972	130	$1.2e^{-4}$	0.966	474	$1.2e^{-4}$	0.962	2008
Rec	$1.2e^{-4}$	0.972	77	$1.2e^{-4}$	0.967	259	$1.1e^{-4}$	0.963	1063
30%	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU	$\overline{ISE}$	$\overline{Cor}$	CPU
N-rec	$2.9e^{-4}$	0.925	130	$2.8e^{-4}$	0.919	498	$2.6e^{-4}$	0.914	2027
Rec	$2.9e^{-4}$	0.927	75	$2.7e^{-4}$	0.920	252	$2.6e^{-4}$	0.916	1060

TABLE 5. Average $ISEs$ and Correlations (approximated using $N = 500$ trials) and total CPU time in seconds of the non-recursive estimator $\widetilde{f}_n$ and the proposed recursive estimator $\widehat{f}_n$ with the choice of the stepsize $(\gamma_n) = (n^{-1})$. Here we consider the chi squared distribution with 12 degrees of freedom $X \sim \chi^2(12)$, with the censoring level equal respectively to 0% (in the first block), equal to 10% (in the second block), equal to 20% (in the third block) and equal to 30% (in the last block), we consider three sample sizes $n = 100$, $n = 200$ and $n = 500$.

and (2.3) follows from remark A.1. In the case $a > \alpha/5$, we have $h_n^2 = o\left(\sqrt{\gamma_n h_n^{-1}}\right)$, and $\lim_{n\to\infty}(n\gamma_n) > (\alpha - a)/2$, then Lemma A.2 ensures that

$$\begin{aligned}
\mathbb{E}[f_n(x)] - f(x) &= Q_n \sum_{k=1}^{n} Q_k^{-1}\gamma_k o\left(\sqrt{\gamma_k h_k^{-1}}\right) + O(Q_n) \\
&= o\left(\sqrt{\gamma_n h_n^{-1}}\right),
\end{aligned}$$

then (2.4) follows from remark A.1. Further, we have

$$\begin{aligned}
Var[f_n(x)] &= Q_n^2 \sum_{k=1}^{n} Q_k^{-2}\gamma_k^2 Var[Z_k(x)] \\
&= Q_n^2 \sum_{k=1}^{n} Q_k^{-2}\gamma_k^2 \left(\mathbb{E}\left(Z_k^2(x)\right) - \left(\mathbb{E}(Z_k(x))\right)^2\right). \qquad (A.8)
\end{aligned}$$

Moreover, in view of (A.6), for $p = 2$, that

$$\begin{aligned}
\mathbb{E}\left(Z_k^2(x)\right) &= h_k^{-1}\pi_k^{-2}\int_{\mathbb{R}} \Delta(x - zh_k) K^2(z)\,dz \\
&= h_k^{-1}\pi_k^{-2}\Delta(x)\int_{\mathbb{R}} K^2(z)\,dz + \nu_k(x), \qquad (A.9)
\end{aligned}$$

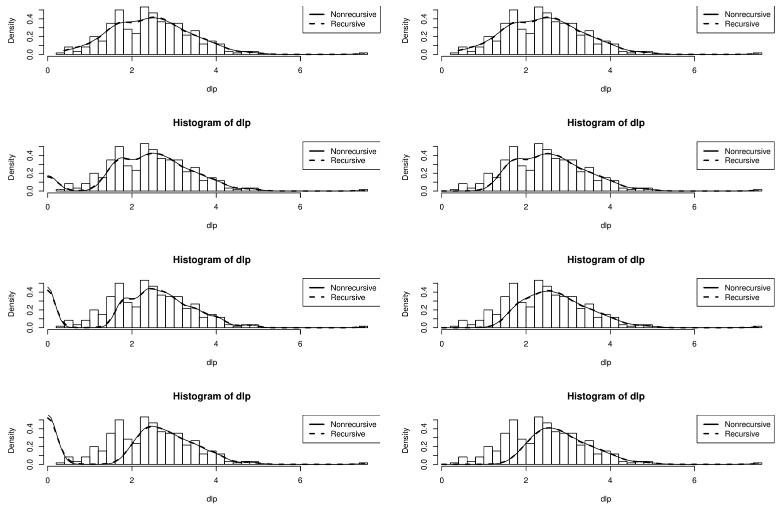

FIGURE 3. Quantitative comparison between the non-recursive estimator (1.4) and the proposed recursive estimator (1.2) with the choice of $(\gamma_n) = \left(n^{-1}\right)$. Here we consider the Parasite Load with the censoring level equal respectively to 0% (in the first block line), equal to 10% (in the second block line), equal to 20% (in the third block line) and equal to 30% (in the last block line), the first block column correspond to the use of the no censored kernel density estimators; the non-recursive estimator introduced by [17] and the recursive estimator given by [10], the second block column correspond to the use of the censored kernel density estimators, we consider three sample sizes $n = 100$, $n = 200$ and $n = 500$, the number of simulations is 500.

with

$$\nu_k(x) = h_k^{-1}\pi_k^{-2}\int_{\mathbb{R}} K^2(z)\left[\Delta(x - zh_k) - \Delta(x)\right]dz.$$

Moreover, it follows from (A.7), that

$$\mathbb{E}\left[Z_k(x)\right] = f(x) + \widetilde{\nu}_k(x), \tag{A.10}$$

with

$$\widetilde{\nu}_k(x) = \pi_k^{-1}\int_{\mathbb{R}} K(z)\left[\Delta(x - zh_k) - \Delta(x)\right]dz.$$

Then, it follows from (A.8), (A.9) and (A.10), that

$$\begin{aligned} Var\left[f_n(x)\right] &= \Delta(x) R(K) Q_n^2 \sum_{k=1}^n Q_k^{-2}\gamma_k^2 h_k^{-1}\pi_k^{-2} + Q_n^2\sum_{k=1}^n Q_k^{-2}\gamma_k^2\nu_k(x) \\ &\quad -\Delta^2(x) Q_n^2\sum_{k=1}^n Q_k^{-2}\gamma_k^2 - 2\Delta(x) Q_n^2\sum_{k=1}^n Q_k^{-2}\gamma_k^2\widetilde{\nu}_k(x) \\ &\quad -Q_n^2\sum_{k=1}^n Q_k^{-2}\gamma_k^2\widetilde{\nu}_k^2(x). \end{aligned}$$

Since Δ is bounded continuous, we have $\lim_{k\to\infty}\nu_k(x)=0$ and $\lim_{k\to\infty}\widetilde{\nu_k}(x)=0$. In the case $a\geq\alpha/5$, we have $\lim_{n\to\infty}(n\gamma_n)>(\alpha-a)/2$, and the application of Lemma A.2 gives

$$Var\left[f_n(x)\right] = \frac{\gamma_n}{h_n}\pi_n^{-1}\left(2-(\alpha-a)\xi\right)^{-1}f(x)R(K)+o\left(\frac{\gamma_n}{h_n}\right),$$

then, (2.5) follows from remark A.1. Now, in the case $a<\alpha/5$, we have $\gamma_nh_n^{-1}=o\left(h_n^4\right)$, and $\lim_{n\to\infty}(n\gamma_n)>2a$, then the application of Lemma A.2 gives

$$\begin{aligned} Var\left[f_n(x)\right] &= Q_n^2\sum_{k=1}^{n}Q_k^{-2}\gamma_k o\left(h_k^4\right) \\ &= o\left(h_n^4\right), \end{aligned}$$

then, (2.6) follows from remark A.1.

A.2. Proof of Proposition 2.5. Following similar steps as the proof of the Proposition 2 of [10], we proof the Proposition 2.5.

A.3. Proof of Theorem 2.3. Let us at first assume that, if $a\geq\alpha/5$, then

$$\begin{aligned} &\sqrt{\gamma_n^{-1}h_n}\left(f_n(x)-\mathbb{E}\left[f_n(x)\right]\right) \\ &\qquad\stackrel{\mathcal{D}}{\to}\mathcal{N}\left(0,\pi_n^{-1}\left(2-(\alpha-a)\xi\right)^{-1}f(x)R(K)\right). \end{aligned} \tag{A.11}$$

In the case when $a>\alpha/5$, Part 1 of Theorem 2.3 follows from the combination of (2.4), (A.11) and remark A.1. In the case when $a=\alpha/5$, Parts 1 and 2 of Theorem 2.3 follow from the combination of (2.3), (A.11) and remark A.1. In the case $a<\alpha/5$, (2.6) implies that

$$h_n^{-2}\left(f_n(x)-\mathbb{E}\left(f_n(x)\right)\right)\stackrel{\mathbb{P}}{\to}0,$$

and the application of (2.3) and remark A.1 gives Part 2 of Theorem 2.3.

We now prove (A.11). In view of (A.3) , we have

$$f_n(x)-\mathbb{E}\left[f_n(x)\right]=Q_n\sum_{k=1}^{n}Q_k^{-1}\gamma_k\left(Z_k(x)-\mathbb{E}\left[Z_k(x)\right]\right).$$

Set

$$Y_k(x) = Q_k^{-1}\gamma_k\left(Z_k(x)-\mathbb{E}\left[Z_k(x)\right]\right). \tag{A.12}$$

The application of Lemma A.2 ensures that

$$\begin{aligned} v_n^2 &= \sum_{k=1}^{n}Var\left(Y_k(x)\right) \\ &= \sum_{k=1}^{n}Q_k^{-2}\gamma_k^2Var\left(Z_k(x)\right) \\ &= \sum_{k=1}^{n}Q_k^{-2}\gamma_k^2h_k^{-1}\left[f(x)\pi_n^{-1}R(K)+o(1)\right] \\ &= Q_n^{-2}\gamma_nh_n^{-1}\left[\left(2-(\alpha-a)\xi\right)^{-1}f(x)\pi_n^{-1}R(K)+o(1)\right]. \end{aligned} \tag{A.13}$$

On the other hand, we have, for all $p > 0$,

$$\mathbb{E}\left[\left|Z_k(x)\right|^{2+p}\right] = O\left(\frac{1}{h_k^{1+p}}\right), \tag{A.14}$$

and, since $\lim_{n\to\infty}(n\gamma_n) > (\alpha - a)/2$, there exists $p > 0$ such that

$$\lim_{n\to\infty}(n\gamma_n) > \frac{1+p}{2+p}(\alpha - a).$$

Applying Lemma $A.2$, we get

$$\begin{aligned}\sum_{k=1}^{n}\mathbb{E}\left[\left|Y_k(x)\right|^{2+p}\right] &= O\left(\sum_{k=1}^{n}Q_k^{-2-p}\gamma_k^{2+p}\mathbb{E}\left[\left|Y_k(x)\right|^{2+p}\right]\right)\\ &= O\left(\sum_{k=1}^{n}\frac{Q_k^{-2-p}\gamma_k^{2+p}}{h_k^{1+p}}\right)\\ &= O\left(\frac{\gamma_n^{1+p}}{Q_n^{2+p}h_n^{1+p}}\right),\end{aligned}$$

and we thus obtain

$$\frac{1}{v_n^{2+p}}\sum_{k=1}^{n}\mathbb{E}\left[\left|Y_k(x)\right|^{2+p}\right] = O\left(\left[\gamma_n h_n^{-1}\right]^{p/2}\right) = o(1).$$

The convergence in (A.11) then follows from the application of Lyapounov's Theorem.

A.4. Proof of Theorem 2.4. Set

$$S_n(x) = \sum_{k=1}^{n}Y_k(x), \qquad \text{and} \qquad s_n = \sum_{k=1}^{n}\gamma_k,$$

where Y_k is defined in (A.12).

- Let us first consider the case $a \geq \alpha/5$. We let $H_n^2 = \Pi_n^2\gamma_n^{-1}h_n$, then

$$\begin{aligned}\ln\left(H_n^{-2}\right) &= -2\ln(\Pi_n) + \ln\left(\prod_{k=1}^{n}\frac{\gamma_{k-1}^{-1}h_{k-1}}{\gamma_i^{-1}h_k}\right)\\ &= (2 - \xi(\alpha - a))\, s_n + o(s_n).\end{aligned} \tag{A.15}$$

  Since $2 - \xi(\alpha - a) > 0$, it follows in particular that $\lim_{n\to+\infty}H_n^{-2} = \infty$. Moreover, since we have $\lim_{n\to+\infty}H_n^2/H_{n-1}^2 = 1$, it follows from (A.13)

$$\lim_{n\to+\infty}H_n^2\sum_{k=1}^{n}Var\left[Y_k(x)\right] = \frac{\pi_n^{-1}f(x)R(K)}{2 - (\alpha - a)\xi}.$$

  Now, in view of (A.14), we have $\mathbb{E}\left[\left|Y_k(x)\right|^3\right] = O\left(Q_k^{-3}\gamma_k^3 h_k^{-2}\right)$, and then, the application of Lemma A.2 ensures that

$$\begin{aligned}n^{-3/2}\sum_{k=1}^{n}\mathbb{E}\left(\left|H_nY_k(x)\right|^3\right) &= O\left(n^{-3/2}H_n^3\sum_{k=1}^{n}Q_k^{-3}\gamma_k^3h_k^{-2}\right)\\ &= o\left(\left[\ln\left(H_n^{-2}\right)\right]^{-1}\right).\end{aligned}$$

Then, the application of Theorem 1 of [9] ensures that, with probability one, the sequence

$$\left(\frac{H_nS_n(x)}{\sqrt{2\ln\ln\left(H_n^{-2}\right)}}\right)=\left(\frac{\sqrt{\gamma_n^{-1}h_n}\left(f_n(x)-\mathbb{E}\left[f_n(x)\right]\right)}{\sqrt{2\ln\ln\left(H_n^{-2}\right)}}\right)$$

is relatively compact and its limit set is the interval

$$\left[-\sqrt{\frac{\pi_n^{-1}f(x)R(K)}{2-(\alpha-a)\xi}},\sqrt{\frac{\pi_n^{-1}f(x)R(K)}{2-(\alpha-a)\xi}}\right]. \tag{A.16}$$

Moreover, it follows from (A.15), that $\lim_{n\to\infty}\ln\ln\left(H_n^{-2}\right)/\ln s_n=1$, and then, with probability one, the sequence

$$\left(\sqrt{\gamma_n^{-1}h_n}\left(f_n(x)-\mathbb{E}\left[f_n(x)\right]\right)/\sqrt{2\ln s_n}\right)$$

is relatively compact, and its limit set is the interval (A.16). Then, the combination of (2.3), (2.4) and remark A.1 concludes the proof Theorem 2.4 in the case $a\geq\alpha/5$.

- Let us now consider the case $a<\alpha/5$.
We set $H_n^{-2}=Q_n^{-2}h_n^4\left(\ln\ln\left(Q_n^{-2}h_n^4\right)\right)^{-1}$, then

$$\begin{aligned}\ln\left(H_n^{-2}h_n^4\right)&=-2\ln\left(Q_n\right)+\ln\left(\prod_{k=1}^n\frac{h_{k-1}^{-4}}{h_k^{-4}}\right)\\&=(2-4a\xi)s_n+o(s_n).\end{aligned} \tag{A.17}$$

Since $2-4a\xi>0$, it follows in particular that $\lim_{n\to+\infty}H_n^{-2}h_n^4=\infty$. Moreover, since we have $\lim_{n\to+\infty}H_n^2/H_{n-1}^2=1$, the application of A.2, ensures that

$$\lim_{n\to+\infty}H_n^2\sum_{k=1}^nVar\left[Y_k(x)\right]=o(1).$$

Moreover, in view of (A.14), it follows from (A.17) and Lemma A.2 that

$$\begin{aligned}&n^{-3/2}\sum_{k=1}^n\mathbb{E}\left(\left|H_nY_k(x)\right|^3\right)\\&=O\left(n^{-3/2}H_n^3h_n^{-6}\left[\ln\ln\left(Q_n^{-2}h_n^4\right)\right]^{3/2}\sum_{k=1}^nQ_k^{-3}\gamma_k^3h_k^{-2}\right)\\&=o\left(\left[\ln\left(H_n^{-2}\right)\right]^{-1}\right).\end{aligned}$$

The application of Theorem 1 of [9] ensures that, with probability one,

$$\begin{aligned}&\lim_{n\to\infty}\frac{H_nS_n(x)}{\sqrt{2\ln\ln\left(H_n^{-2}\right)}}\\&=\lim_{n\to\infty}h_n^{-2}\sqrt{\frac{\ln\ln\left(Q_n^{-2}h_n^4\right)}{2\ln\ln\left(H_n^{-2}\right)}}\left(f_n(x)-\mathbb{E}\left[f_n(x)\right]\right)=0.\end{aligned}$$

Moreover, since (A.17) ensures that $\lim_{n\to\infty} \ln\ln\left(H_n^{-2}\right)/\ln\ln\left(Q_n^{-2}h_n^4\right) = 1$, we obtain

$$\lim_{n\to\infty} h_n^{-2}\left(f_n(x) - \mathbb{E}\left[f_n(x)\right]\right) = 0 \quad \text{a.s.},$$

then, the second part of Theorem 2.4 follows from (2.3) and remark A.1.

A.5. Proof of Corollary 2.10. Following similar steps as in [20], we prove first the following corollary

Corollary A.3. *Let the assumptions* $(A1)-(A3)$ *hold, and the bandwidth* (h_n) *equal to (2.12) and the stepsize* $(\gamma_n) = \left(n^{-1}\right)$ *when we apply the estimator* $\widehat{f}_n$ *and the bandwidth* (h_n) *equal to (2.17) when we apply the estimator* $\widetilde{f}_n$. *We have*

$$\begin{aligned}
&\mathbb{E}\left[\widehat{MWISE}\left[\widehat{f}_n\right]\right] \\
&\quad= \frac{5}{4}2^{2/5}\left(\frac{5}{6}\right)^{6/5}\Theta(K)\, I_1^{4/5}I_2^{1/5} \\
&\quad\left[\left(1+0.26522\times C_1 I_2^{-1}n^{-3/7}\right.\right. \\
&\quad\left.+\left(0.08793\times C_2 + 0.03521\times C_3\right)I_2^{-1}n^{-6/7} - 0.22316\times n^{-1}\right) \\
&\quad\left.+o\left(n^{-1}\right)\right]n^{-4/5}\left(1+o(1)\right)
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\left[\widetilde{MWISE}\left[\widetilde{f}_n\right]\right] \\
&\quad= \frac{5}{4}\Theta(K)\, I_1^{4/5}I_2^{1/5} \\
&\quad\left[\left(1+0.22797\times C_1 I_2^{-1}n^{-3/7}\right.\right. \\
&\quad\left.\left.+\left(0.06496\times C_2 + 0.02165\times C_3\right)I_2^{-1}n^{-6/7}\right)\right]n^{-4/5}\left(1+o(1)\right),
\end{aligned}$$

where C_1, C_2 and C_3 are some quantities depending on the density f and on Δ. Then, we can deduce that for some specific case and for n small enough the expectation of $\widehat{MWISE}\left[\widehat{f}_n\right]$ can be smaller than the expectation of $\widetilde{MWISE}\left[\widetilde{f}_n\right]$.

Acknowledgments

The author would like to thank the Editor and the referees for their helpful comments.

References

1. Blum, J. R. and Susarla, V.: Maximal deviation theory of density and failure rate function estimates based on censored data. *North-Holland, Amsterdam-New York*, 1980.
2. Bojanic, R. and Seneta, E.: A unified theory of regularly varying sequences. *Math. Z.*, **134** (1973), no. 2, 91–106.
3. Diehl, S. and Stute, W.: Kernel density and hazard function estimation in the presence of censoring. *J. Multivariate Anal.*, **25** (1988), no. 2, 299–310.

4. Földes, A., Rejtö, L. and Winter, B. B.: Strong consistency properties of nonparametric estimators for randomly censored data, part II: Estimation of density and failure rate. *Period. Math. Hungar*, **12** (1981), no. 1, 15–29.
5. Galambos, J. and Seneta, E.: Regularly varying sequences. *Proc. Amer. Math. Soc.*, **41** (1973), no. 1, 110–116.
6. Jmaei, A., Slaoui, Y., and Dellagi, W.: Recursive distribution estimators defined by stochastic approximation method using Bernstein polynomials. *J. Nonparametr. Stat.*, **29** (2017), no. 4, 792–805.
7. Milet, J., Nuel, G., Watier, L., Courtin, D., Slaoui, Y., Senghor, P., Migot-Nabias, F., Gaye, O. and Garcia, A.: Genome Wide Linkage Study, Using a 250K SNP Map, of Plasmodium falciparum Infection and Mild Malaria Attack in a Senegalese Population. *PLoS ONE*, **5** (2010), no. 7, e11616.
8. Mokkadem, A. and Pelletier, M.: A companion for the Kiefer-Wolfowitz-Blum stochastic approximation algorithm. *Ann. Statist.*, **35** (2007), no. 4, 1749–1772.
9. Mokkadem, A. and Pelletier, M.: Compact law of the iterated logarithm for matrix-normalized sums of random vectors. *Theory Probab. Appl.*, **52** (2007), no. 4, 2459–2478.
10. Mokkadem, A, Pelletier, M. and Slaoui, Y.: The stochastic approximation method for the estimation of a multivariate probability density. *J. Statist. Plann. Inference*, **139** (2009), no. 7, 2459–2478.
11. Nadaraya, E. A.: On estimating regression. *Theory Probab. Appl.*, **9** (1964), no. 1, 141–142.
12. Parzen, E.: On estimation of a probability density and mode. *Ann. Math. Statist.*, **33** (1962), no. 3, 1065–1076.
13. Révész, P.: Robbins-Monro procedure in a Hilbert space and its application in the theory of learning processes I. *Studia Sci. Math. Hung.*, **8** (1973), no. 1, 391–398.
14. Révész, P.: How to apply the method of stochastic approximation in the non-parametric estimation of a regression function. *Math. Operationsforsch. Statist., Ser. Statistics.*, **8** (1977), no. 1, 119–126.
15. Robbins, H. and Monro, S.: A stochastic approximation method. *Ann. Math. Stat.*, **22** (1951), no. 2, 400–407.
16. Rosenbaum, P. R and Rubin D. B.: The central role of the propensity score in observational studies for causal effects. *Biometrika*, **70** (1983), no. 1, 41–55.
17. Rosenblatt, M.: Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.*, **27** (1956), no. 3, 832–837.
18. Silverman B. W.: Density estimation for statistics and data analysis. *Chapman and Hall*, London, 1986.
19. Slaoui, Y.: Large and moderate principles for recursive kernel density estimators defined by stochastic approximation method. *Serdica Math. J.*, **39** (2013), no. 1, 53–82.
20. Slaoui, Y.: Bandwidth selection for recursive kernel density estimators defined by stochastic approximation method. *J. Probab. Stat*, (2014), ID 739640, doi:10.1155/2014/739640.
21. Slaoui, Y. : The stochastic approximation method for the estimation of a distribution function. *Math. Methods Statist.*, **23** (2014), no. 4, 306–325.
22. Slaoui, Y.: Plug-In Bandwidth selector for recursive kernel regression estimators defined by stochastic approximation method. *Stat. Neerl.*, **69** (2015), no. 4, 483–509.
23. Slaoui, Y.: Optimal bandwidth selection for semi-recursive kernel regression estimators. *Stat. Interface*, **9** (2016), no. 3, 375–388.
24. Slaoui, Y.: Recursive kernel density estimators under missing data. *Comm. Statist. Theory Methods*, **46** (2017), no. 18, 9101–9125.
25. Slaoui, Y. and Jmaei, A.: Recursive density estimators based on Robbins-Monro's scheme and using Bernstein polynomials. *Stat. Interface*, **12** (2019), no. 3, 439–455.
26. Slaoui, Y. and Nuel, G.: Parameter estimation in a hierarchical random intercept model with censored response: An approach using a SEM algorithm and Gibbs sampling. *Sankhya B*, **76** (2014), no. 2, 210–233.
27. Tsybakov, A. B.: Recurrent estimation of the mode of a multidimensional distribution. *Probl. Inf. Transm.*, **26** (1990), no. 1, 31–37.
28. Watson, G. S.: Smooth regression analysis. *Sankhya A*, **26** (1964), no. 4, 359–372.

Yousri Slaoui: Univ. Poitiers, Lab. Math. et Appl., Futuroscope Chasseneuil, France

E-mail address: Yousri.Slaoui@math.univ-poitiers.fr

URL: http://www-math.sp2mi.univ-poitiers.fr/~yslaoui/